



⑪ Publication number : **0 598 512 A1**

⑫

## EUROPEAN PATENT APPLICATION

⑫① Application number : **93308627.4**

⑫⑤ Int. Cl.<sup>5</sup> : **G06F 3/14**

⑫② Date of filing : **28.10.93**

⑫③ Priority : **30.10.92 GB 9222767**

⑫④ Date of publication of application :  
**25.05.94 Bulletin 94/21**

⑫⑧ Designated Contracting States :  
**DE ES FR GB IT NL**

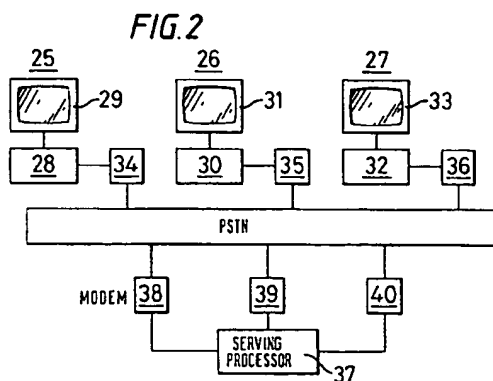
⑫⑦ Applicant : **CANON KABUSHIKI KAISHA**  
**30-2, 3-chome, Shimomaruko,**  
**Ohta-ku**  
**Tokyo (JP)**

⑫⑦ Inventor : **Chudley, Martin John c/o CANON**  
**Res. Cent. Europe**  
**17-20 Frederic Sanger Rd**  
**The Surrey Research Park**  
**Guildford, Surrey GU2 5YD (GB)**

⑫④ Representative : **Beresford, Keith Denis Lewis**  
**et al**  
**BERESFORD & Co.**  
**2-5 Warwick Court**  
**High Holborn**  
**London WC1R 5DJ (GB)**

⑫④ Processing image data.

⑫⑦ A plurality of stations (25, 26, 27) are connected to a central serving processor (37) via the public switched telephone network (33). Each station (25, 26, 27) includes a processing device (28) and a monitor (29) for generating an interactive three-dimensional graphics environment. Similar environments are provided on each remote station and interaction between the stations is facilitated by transmitting transformation matrices, modified by one station, to all of the other stations. Each transformation matrix defines a transformation of an object or a light source into the displayable space.



The present invention relates to apparatus and methods for processing image data.

In particular, the present invention relates to apparatus and methods for displaying three-dimensional images. Furthermore, the present invention relates to an interactive graphics system in which a plurality of users may access three-dimensional data.

Systems are known which are capable of synthesizing two dimensional images in response to data defining elements within a three dimensional space. The final two-dimensional result may consist of a very large number of colored pictured elements (pixels) which may be viewed on a monitor or printed onto an image carrying medium.

In interactive systems arranged to generate data representing a three-dimensional space, objects appear to move within the three-dimensional space in response to input commands. Thus, in such systems a machine is required to render a two-dimensional image from data representing a three-dimensional space repeatedly, as the position and/or orientation of objects, light sources and view position change in response to input commands. Typically, in an interactive environment, a machine is required to produce output images at a rate of between five to twenty per second.

Interactive three-dimensional graphics systems have many applications and these applications could be extended further if many users could share access to a common three-dimensional world. Thus, information could be recorded in three-dimensional space and many users, each with their own display device, could be given access to this space, allowing them to observe the data created by other users and, possibly, modify this data. Thus, an interchange of ideas could be provided by giving access to shared three-dimensional data.

In interactive graphics systems output images are usually displayed on a monitor, similar to a conventional television receiver. Thus, a processor is required to generate a video signal which is supplied to the monitor at video rate. Often data is supplied to a frame buffer at an interactive rate and image data is then read from the frame buffer at video rate, in order to generate the video signal. Typically, a look-up table is provided, allowing a selection of colors to be defined in response to selections made by the limited number of bits stored in the look-up table. The look-up table itself may define particular colors and include values representing modifications to luminance and saturation.

Techniques for the transmission of video signals are well known and, in addition to one-way television broadcasts, it is also known to provide two-way video communication, such as in video conferencing facilities or video telephones.

However, a problem with transmitting video sig-

nals is that a substantial bandwidth is required, therefore transmission costs become high. A further problem arises, when sharing access to a three-dimensional world, in that additional communication channels are required for the control of access to that world, along with processing facilities to resolve contention problems. Furthermore, although the video image includes a large amount of data, it only represents a particular view of the three-dimensional world, rendered from a particular view point. Thus, any attempt to provide multiple access to the three-dimensional data is restricted to the particular view under consideration.

According to a first aspect of the present invention, there is provided apparatus for displaying three-dimensional images, comprising a first station having a processing device and a display device, and a second station having a processing device and a display device, wherein modifications made within a three-dimensional space displayable by said first station result in similar modifications to similar space displayable by said second station, each station including means for generating a video image from three-dimensional data and, on effecting a modification within the space displayable by said first station, means for transmitting modified three-dimensional data to said second station over a communication channel.

Thus, the present invention solves the problem of providing shared access by transmitting modified three-dimensional data between stations.

Three-dimensional data may represent shape geometry and transformations for transforming said geometry between frames of reference, such as between a modelling space for each specific object, three-dimensional world space and viewing space. In a preferred embodiment, data defining the transformations of objects are transmitted between stations.

Possibly, data representing transformation differences could be transmitted but in a preferred embodiment data representing a complete transformation is transmitted after said data has been modified. Preferably, the transformation data represents matrices.

In a preferred embodiment a central serving processor is provided and a plurality of users are allowed to modify data under the control of said server.

The invention will now be described by way of example only, with reference to the accompanying drawings, in which:

Figure 1 shows an interactive three-dimensional graphics station, including a processor and a modem;

Figure 2 shows a plurality of client machine stations, of the type shown in Figure 1, connected to a server machine;

Figure 3 is an overview of the operations performed by a client machine of the type shown in

Figure 2, including communication procedures and rendering procedures;

Figure 4 details the rendering procedures identified in Figure 3;

Figure 5 details the communication procedures identified in Figure 3; and

Figure 6 details communication procedures for the server machine identified in Figure 2.

A system for processing image data representing three-dimensional objects is shown in Figure 1. A processor 15 is arranged to write data to and read data from an internal memory device. Data stored in the memory device defines three-dimensional image data, two-dimensional image data and instructions to the processor, including instructions relating to communications with a central server.

The processor 15 receives input data from input devices, consisting of a manually operable keyboard 17 and a position sensitive device, such as a trackball 18, a mouse or a digitising stylus etc.

Two-dimensional images are displayed on a visual display unit 19, which receives output image data at video rate by raster scanning a frame buffer, often provided with proprietary display units. The visual display unit 19 has a definition of, typically, one thousand lines with one thousand pixels on each line, requiring a frame buffer with one million pixel locations. For the bulk transfer of program data and image data, a mass storage device 20 is provided, such as a hard magnetic disk, optical disk or tape drive etc.

The processor 15 is connected to a modem 21, which is in turn connected to a conventional public telephone line. Thus, modem 21 allows processor 15 to communicate over any distance, using the public switched telephone network. The processor 15 may include instructions for initiating a connection to another system or, alternatively, means for establishing such a connection with another modem may be provided within modem 21.

A system embodying the present invention is shown in Figure 2, in which, in this example, three systems of the type shown in Figure 1 are provided. A first station 25 includes a first processing device 28 and a first display device 29. Similarly, a second station 26 includes a second processing device 30 and a second display device 31, while a third processing station includes a third processing device 32 and a third display device 33. Each processing device 28, 30, 32 is connected to the public switched telephone network (PSTN) 33 via respective modems 34, 35, 36. The PSTN 33 represents the global switched telecommunications network, allowing international communication. Thus, station 29 could be in Munich, station 26 in Washington and station 27 in Tokyo. Alternatively, other systems may be employed, such as dedicated lines or higher capacity switched networks, such as a packet switching network.

Each processor 28, 30, 32 of the user stations 25,

26, 27 communicates with a central serving processor 37 via respective modems 38, 39, 40. Thus, direct communication between remote stations is not permitted and all communications over the network are controlled by the serving processor 37.

The serving processor 37 is shown without an associated display device. In some situations a display device may be provided, although it is not essential to the operation of the system. It is envisaged that actual use of the networked environment is made by the remote stations.

However, the server does maintain a complete record of all three-dimensional data, so that said data may be supplied over the communication network when, for example, a new user station is connected to said network.

Three remote stations are shown in Figure 2; it is also envisaged that many more stations could be connected to the network and the extent to which said stations may be given access to the graphics environment may vary. Thus, some stations may be allowed to make modifications while other stations may only be permitted to view what has already been created by others.

The processes performed by the processor 15, for creating an interactive three-dimensional graphics environment, are shown in Figure 3. These processes are performed by processors 28, 30 and 32 as shown in Figure 2, identified as client machines, as distinct from the serving machine 37 in Figure 2.

At step 41, communications procedures are adopted which allow the processor to communicate with the serving processor via the PSTN 33 and their respective modems. The operations performed at step 41 are detailed in Figure 5.

At step 42, a question is asked as to whether it is necessary to update the display, that is to say, whether it is necessary to render the three-dimensional data in order to generate new two-dimensional data. If this question is answered in the affirmative, step 43 executes the rendering process, detailed with reference to Figure 4. Thereafter, control is returned to step 41, allowing further communications to be made. Similarly, if the question raised at step 42 is answered in the negative, control is returned to step 41.

Figure 4 details the rendering operation identified as step 43 in Figure 3. At step 45 object culling is performed by determining, after local transformation and viewing transformation, the extent to which an object can actually be seen. Extents are defined for the object in each dimension, that is to say, maximum and minimum values are determined for the object in each of the X, Y and Z dimensions. This process is facilitated by defining a cuboid bounding box of planes of constant x, y or z extent. The bounding box is transformed into viewing space and the position of the transformed bounding box is considered, to deter-

mine whether it is totally within the viewable space, totally outside the viewable space or intersecting the viewable space.

If an object is totally outside the viewable space no further processing is performed on the polygons defining such an object, thereby making a significant saving in processing time. For objects totally within the viewing space, processing continues on these polygons and no saving is made, however, the processing overhead for transforming the bounding box is quite modest. For objects which are partially within the viewable space and partially outside the viewable space, data defining the planes of intersection are stored for reconsideration at step 49, involving the clipping of polygons.

A clock is present which effectively gives a unique time stamp to each iteration. This stamp is then applied each time lighting is moved and each time surfaces are moved. Thus, if on a present iteration, it can be seen that no changes have occurred to the position or orientation of the lighting, or to the position of polygons, from the last iteration, lighting calculations, calculated on the previous iteration, are used again.

At step 47 the lighting of polygons, where necessary, is performed. In the system shown in Figure 3, a local transformation is performed at step 21, surfaces are lit at step 22 and a view transformation is effected at step 23. The local transformation and the view transformation are mathematically similar and, when performed by matrix multiplication, the matrices defining the transformations are concatenated into a common matrix. Thus, the local transformation and the viewing transformation are performed as one transformation from object space to viewing space. The viewing space represents a particular view of the total displayable space.

Before the viewing transformation can be effected, it is necessary to perform lighting calculations. In the present embodiment, this is effected by calculating the inverse to the local transformation, effecting this inverse transformation upon the position of light sources, thereby effectively transforming the light sources into object space. Thereafter, lighting calculations are performed in object space and the object is transformed directly from object space to viewing space, using the concatenated matrix. Thus, the concatenated matrix used to transform the bounding box during the object cull, at step 45, is used again to transform the actual object polygons from object space into viewing space, at step 48.

Step 49 involves the clipping of polygons which do not lie within the viewing space. As previously stated, additional information is generated at step 45, during the object culling process, which identifies planes bounding the viewing space through which objects intersect.

At step 50, the three-dimensional object is pro-

jected into a two-dimensional image and at step 52, backface culling is performed which, in the present embodiment, does not require reference to the polygon normals. Step 52 provides scan conversion and during this process account is taken of the restricted size of most frame buffers which typically have only eight bits per pixel location. Thereafter, a two-dimensional video image is produced by raster scanning the frame buffer at video rate.

Thus, it can be appreciated that the three-dimensional image is constructed by transforming object geometry into a viewing space, from which a particular two-dimensional view is rendered. The position of object geometry, lighting and view position is defined by matrices and, during interactive operation, it will be the values within these matrices which undergo modification.

In the preferred embodiment, transformations are defined by matrices having four rows and four columns with four bytes of data allocated to each position within the matrix, facilitating floating point manipulations. Each matrix therefore consists of 64 bytes of data.

In the preferred embodiment, each remote station 25, 26 and 27 has its own local representation of image geometry. The geometry may be the same, such that substantially the same image is seen at each station. Alternatively, geometry may differ substantially or alternatively and perhaps more appropriately, the geometry may differ only in more subtle ways. Thus, for example, each station may include geometry representing a chair and said chair may have the same unique identification name at each station. However, the chair may be different in design or may be a different color at each of the different stations.

As previously stated, interactive operation at a station results in modification to transformation matrices. The system shown in Figure 2 operates by transmitting new matrix data to all of the stations, such that an interactive modification in three-dimensional space at one station will result in a similar modification to the space at all of the other stations. Thus, each station maintains its own version of a shared three-dimensional world and the amount of data required to maintain equivalent versions consists only of the matrices required to effect transformations. Thus, even when connected by a channel having very limited bandwidth, such as a telephone line, each station is updated with interactive modifications made at any of the other stations.

Input commands from the keyboard 17 or the tracker ball 18, are arranged to provide information to the communications process 41, identifying an operator's desire to lock an object, unlock an object, modify the position of an object, delete an object or create an object. The rendered output displayed to each client is dependent upon that client's own view position

within the three-dimensional space, identified within said space by a suitable object such as a camera. Other cameras are present within the space, identifying the view position of other clients within the network and each camera remains permanently locked to its respective client, so that the particular client cannot move the cameras of others and the client's own camera cannot be moved by the other clients.

Before an object may be modified by a client, it is necessary for that object to be locked and the locking and unlocking of objects is controlled by the central server. Thus, a client may request access to an object, which may or may not be given by the server, dependent upon whether any other client has previously locked that particular object. During operation, a client may lock as many objects as are available and when locked, access to these objects is not permitted by any of the other clients in the network.

Referring to Figure 3, on entering step 41, enquiries are made as to whether any commands have been generated by the client operator. Referring to Figure 5, a question is asked at step 51 as to whether a request has been made to lock an object. If this question is answered in the negative, a question is asked at step 52 as to whether a command has been made to unlock an object. If this question is answered in the negative, a further question is asked at step 53 as to whether the operator wishes to modify an object. Again, if this question is answered in the negative, a further question is asked at step 54 as to whether the operator wishes to delete an object and if this question is answered in the negative, a question is asked at step 55 enquiring as to whether the operator wishes to create an object. If this final question at step 55 is answered in the negative, all possible meaningful input commands have been tested, thereby completing the communications iteration and passing control to step 42 in Figure 3, identified in Figure 5 by the arrow returning control to the top of the loop.

If the question asked at step 51 is answered in the affirmative, a question is asked at step 56 as to whether the object has already been locked. When an object has been locked by a client, this information is supplied to all other clients in the network, therefore a list of locked objects is retained by all clients, allowing the question asked at step 56 to be answered. If the question asked at step 56 is answered in the affirmative, to the effect that the object has been locked, the object is not available to the client and the procedure terminates.

If the object has not been locked and the question asked at step 56 is answered in the negative, a lock request is made at step 57 over the communications channel. In response to this lock request, the central server issues a lock message, thus a question is asked at step 58 as to whether the object has been locked and this question is only answered in the affirmative when the lock message from the server has

been received. The receipt of this message allows the processor to leave the loop surrounding step 58, allowing a further iteration to be made.

It should be noted that, during these communication procedures, no modifications will have been made to the graphics display, therefore the question asked at step 42 will be answered in the negative and control will be returned to the process shown in Figure 5. Thus, after requesting the locking of an object, a further instruction may be made to modify the position of the object. Such an instruction results in the question asked at step 53 being answered in the affirmative and a further question being asked at step 59 as to whether the object to be modified has been locked by the client wishing to modify it. If the question asked at step 59 is answered in the negative, modification is not permitted and control is returned. However, if the client wishing to modify an object has successfully locked use of that object, data relating to the modifications are supplied to the central server at step 60.

After making modifications, the client would wish to unlock the object so that other clients may gain access to it. Again, an input command is made to this effect which, on the next iteration, is identified by the question asked at step 52 being answered in the affirmative. This results in a question being asked at step 61 as to whether the object was previously locked by this client. Again, if this question is answered in the negative, control is returned but if this question is answered in the affirmative, an instruction is sent to the server to the effect that the object is to be unlocked. Again, this information is relayed by the server to all other clients within the network.

The question asked at step 54 allows a client to delete an object and again the question is asked at step 63 as to whether the object has previously been locked by the client wishing to make the request to delete. If this question is answered in the affirmative, deletion instructions are generated at step 64, which are supplied to the central server, thereby deleting the object from the space under consideration. Again, the server will supply this information to all other clients.

A similar procedure is provided in response to the question asked at step 55 which, when answered in the affirmative, allows an object to be created at step 65.

Operation of the serving processor 37 is detailed in Figure 6. In a preferred embodiment, server 37 is only concerned with the communication of data between the clients and does not itself have to perform a rendering operation.

Not being required to perform rendering allows the serving processor 37 to wait for external communications at step 71. Upon receiving information from a client at step 72, enquiries are made as to the nature of the request, at steps 73, 74, 75, 76 and 77. If all of

these questions are answered in the negative, the input data is assumed to be in error and control is returned to the waiting condition at step 71.

A request to lock an object is identified at step 73 and results in a question being asked at step 78 as to whether the object has already been locked. This question should be answered in the negative, given that a request to lock a previously locked object would have been rejected at step 56 in Figure 5, thereby preventing a lock request from being sent. On answering the question raised at step 78 in the negative, a command is sent back to the client machine at step 79, confirming the lock and control is returned to the wait condition at step 71.

A request to modify an object is identified at step 75 and results in a question being asked at step 80 as to whether the object was previously locked by the requesting client. Again, this question would normally be answered in the affirmative, given that a similar check is made at step 59. On the question at step 80 being answered in the affirmative, communication is permitted at step 81, allowing the modifications to be effected.

Requests to unlock an object are identified at step 74 and again a check is made at step 82 as to whether the object was previously locked by the client requesting an unlock. Again this question should be answered in the affirmative, given that a similar check is raised at step 61 and thereafter an unlock command is generated at step 83.

Deletion requests are identified at step 76 and again a question is asked at step 84 as to whether the client requesting deletion has previously locked use of the object. On the question raised at step 84 being answered in the affirmative, communication relating to the deletion is permitted at step 85.

Object creation request are identified at step 77 and communication relating to the creation of objects is carried out at step 86.

Typically, modems 34, 35, 36, 38, 39, 40 provide transmission at 1600 bytes per second, which allows one selected object to be updated twenty times per second. Time stamping information is added to each update, thereby ensuring that updates are re-sequenced, if necessary at a receiver into the correct order. Thus, the correct interval between each update can be re-established at the receiving station independently of the transmission speed, thereby faithfully reproducing an original movement.

It is important that communications requesting access to an object and communications relinquish the locking of an object are conveyed accurately using reliable codes, such as TCP-IP. However, given that a complete matrix defining a complete transformation is transmitted, it is allowable for data relating to matrices to be lost, given that the correct data will be sent on the next iteration. Furthermore, it is possible for terminals and communication channels to

operate at different speeds, wherein some terminals will receive all iterations while others receive a lower number of samples.

## Claims

1. Apparatus for displaying three-dimensional images, comprising
  - a first station (25) having a processing device (15) and a display device (19) and
  - a second station (26) having a processing device and a display device, wherein
  - modifications made within a three-dimensional space displayable by said first station result in similar modifications to similar space displayable by said second station, characterised in that
  - each station including means for generating a video image from three-dimensional data and, on effecting a modification within the space displayable by said first station, means for transmitting modified three-dimensional data to said second station over a communication channel.
2. Apparatus according to claim 1, wherein said three-dimensional data includes geometry data defining objects and transformation data defining the position of said objects, wherein said transformation related data is transmitted.
3. Apparatus for displaying three-dimensional images according to claim 2, wherein complete new transformation data is transmitted for a selected object or light source.
4. Apparatus according to claim 2, wherein said transformation data represents matrices.
5. Apparatus according to claim 1, wherein said communication channel forms part of a telephone network (PSTN).
6. Apparatus according to claim 1, wherein a plurality of remote stations (25, 26, 27) are connected by a central server (37), wherein said server controls the access permitted by said remotest users.
7. A method of displaying three-dimensional images, wherein modifications within a three-dimensional space displayable by a first station result in similar modifications to similar space displayable by a second station, each station being arranged to generate a video image from three-dimensional data and, on effecting a modification within the displayable space, to transmit modified three-dimensional data to said second

station over a communication channel.

8. A method according to claim 7, wherein said three-dimensional data includes geometry data defining objects and transformation data defining the position of objects, wherein said transformation related data is transmitted. 5
9. A method according to claim 8, wherein complete new transformation data is transmitted for a selected object or light source. 10
10. A method according to claim 8, wherein said transformation data represents matrices. 15
11. A method according to any of claims 7 to 10, wherein the communication channel forms part of a telephone network.
12. A method according to any of claims 7 to 11, wherein a plurality of remote stations are connected by a central server and said server controls the access permitted by remote users. 20

25

30

35

40

45

50

55

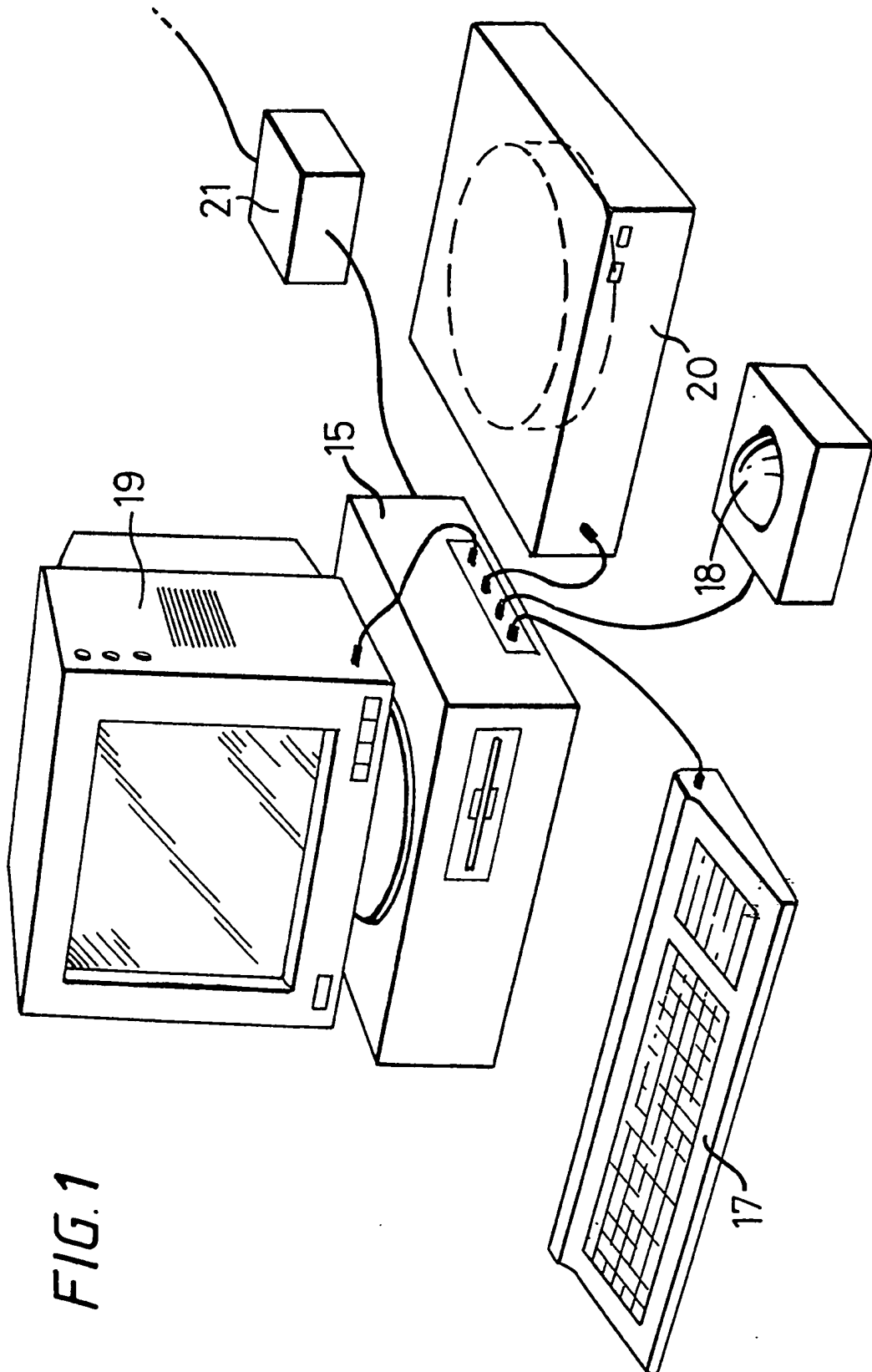
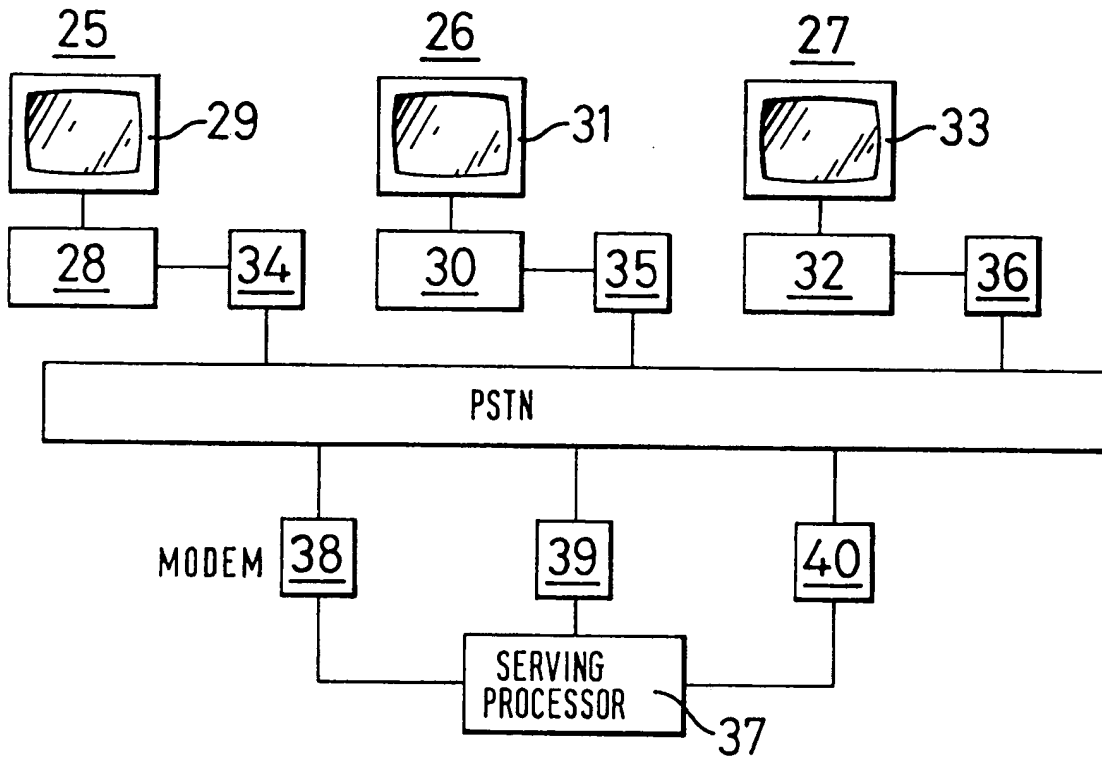


FIG. 1



**FIG. 2**



**FIG. 3** OVERALL FLOW OF CONTROL IN CLIENT.

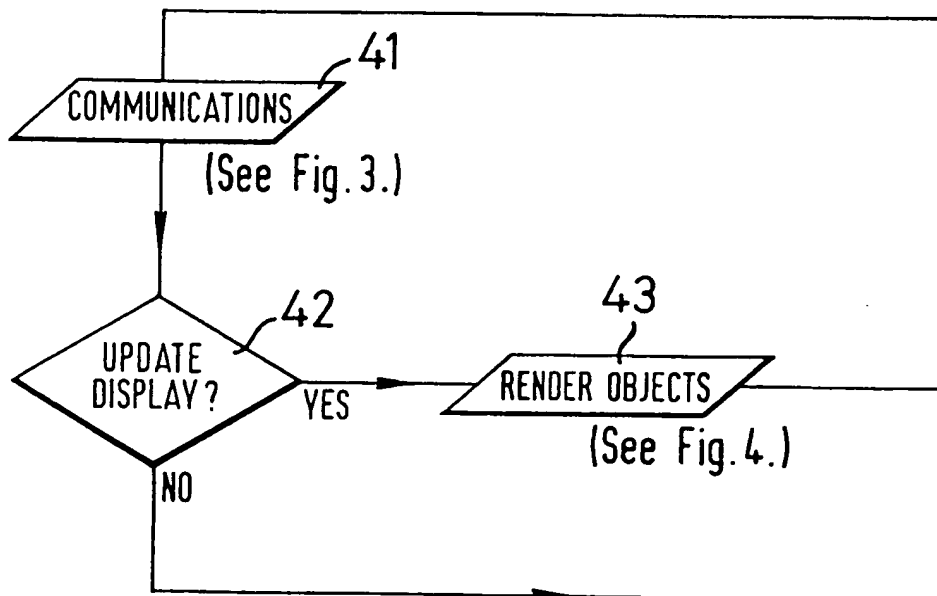
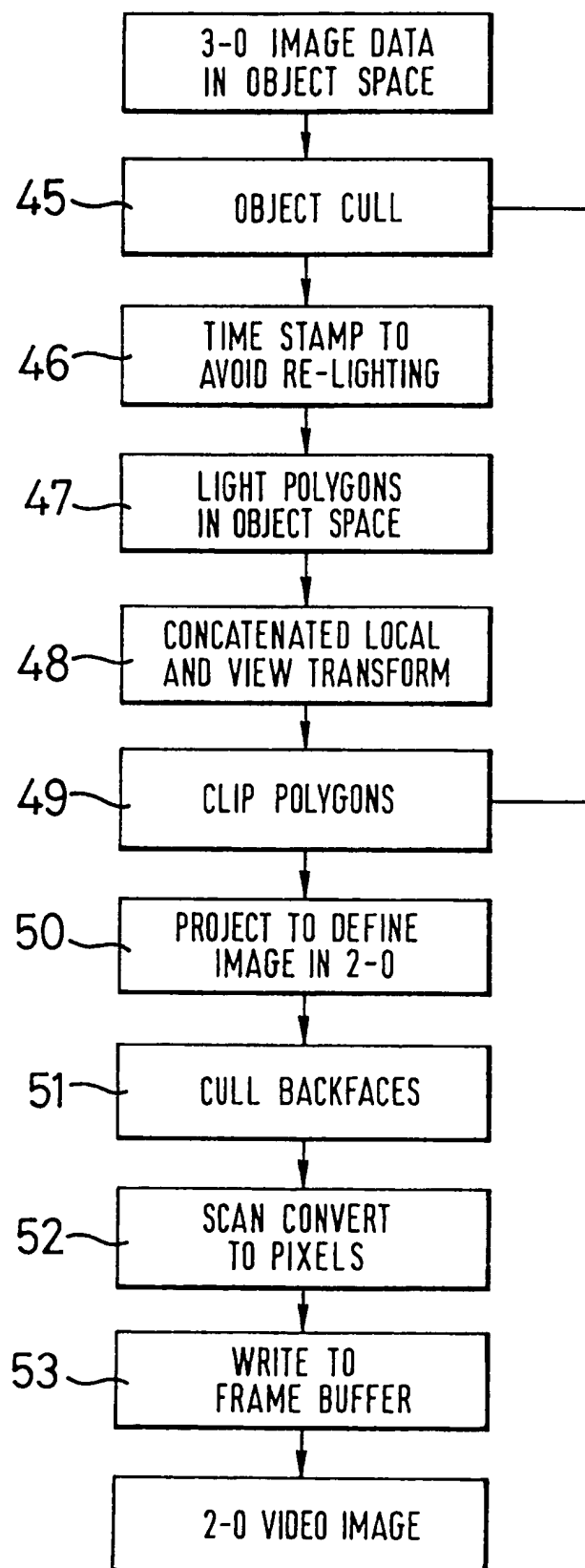
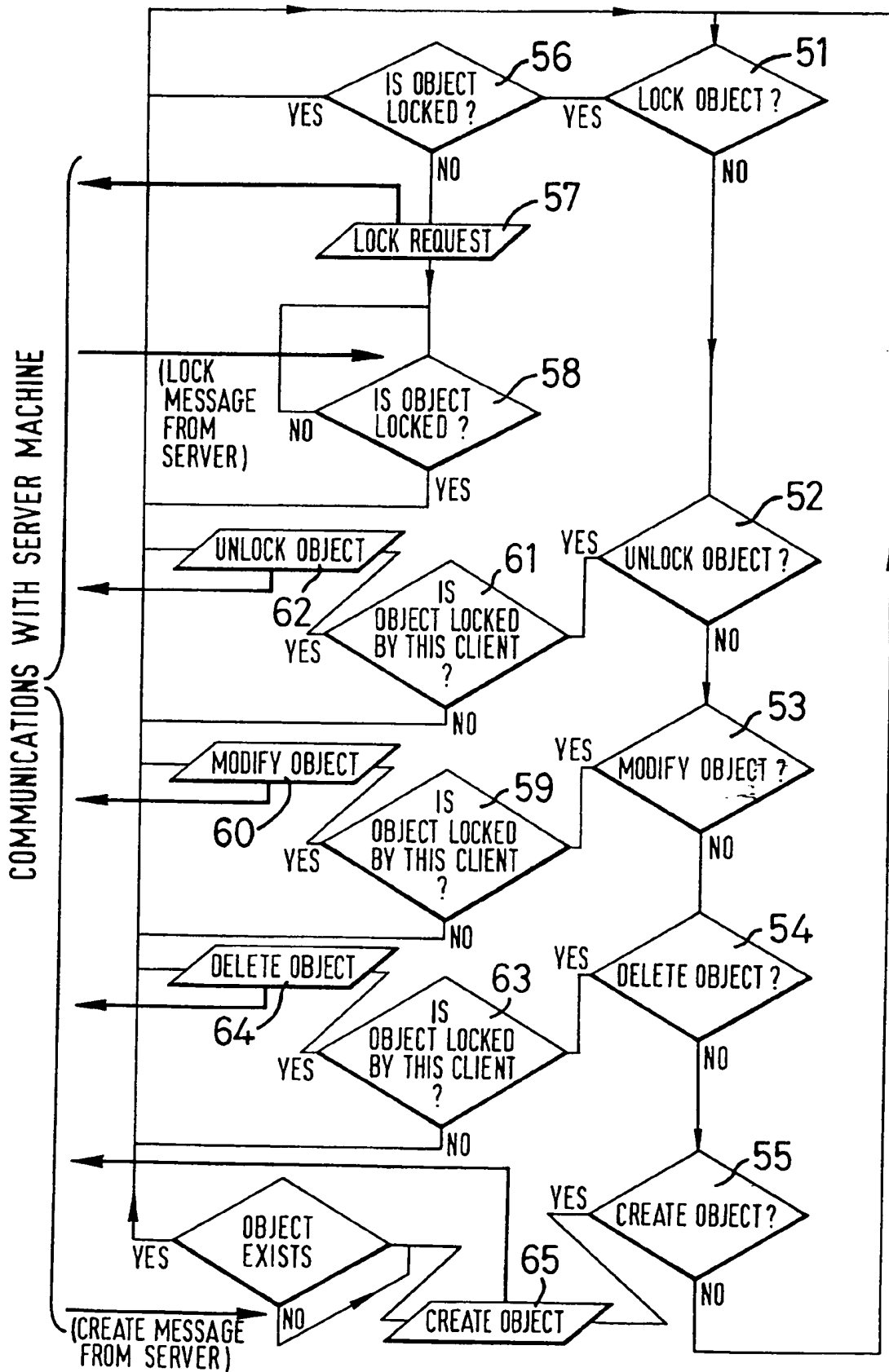
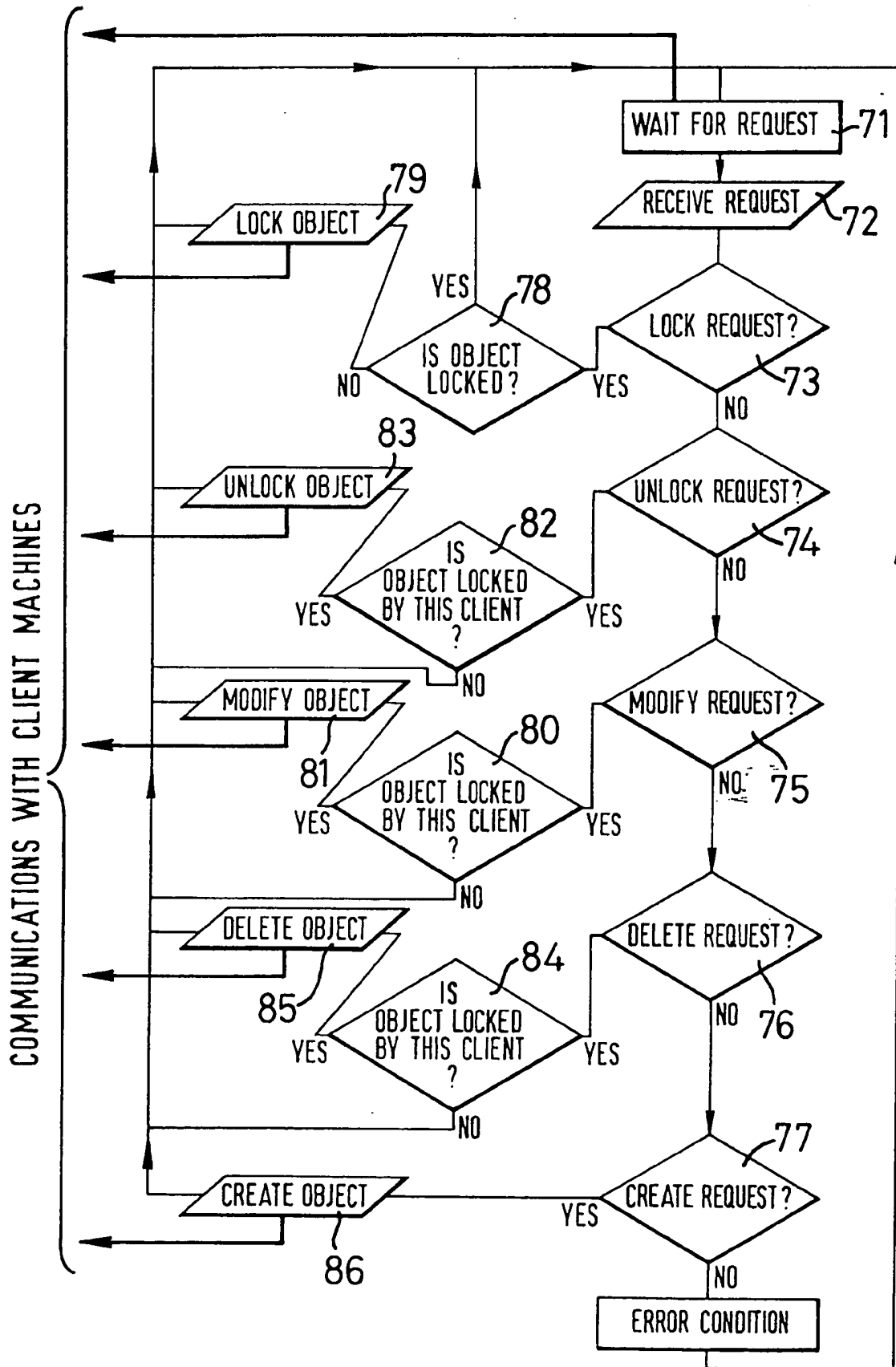


FIG. 4



**FIG.5** OPERATION OF COMMUNICATION IN CLIENT MACHINE.



**FIG. 6** OPERATION OF SERVER MACHINE.



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 93 30 8627

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.5)
A	EP-A-0 475 581 (HEWLETT-PACKARD CO.)  * page 3, column 1, line 35 - column 2, line 13 * * page 4, column 3, line 41 - column 4, line 1 * * page 4, column 4, line 55 - page 5, column 5, line 22 * * claim 1; figures 1,2 * ---	1,7,13,14	G06F3/14
A	US-A-4 414 621 (BOWN ET AL.) * abstract * * figure 1 * ---	5,11	
A	PROCEEDINGS OF IEEE CONFERENCE ON COMMUNICATIONS SOFTWARE: COMMUNICATIONS FOR DISTRIBUTED APPLICATIONS & SYSTEMS. TRICOMM '91 18 April 1991, NEW YORK, N.Y. USA pages 145 - 157 P.F. FITZGERALD ET AL. 'Evaluating Alternative Display Sharing System Architectures' * whole article * -----	1,6,7,12	
			TECHNICAL FIELDS SEARCHED (Int.Cl.5)  G06F
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 14 February 1994	Examiner Farricella, L
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ----- & : member of the same patent family, corresponding document	

EPO FORM 1500 (12.92) (P04C01)